No. 2003–110

# BOB RUSSELL VOLUME
# DON'T AGGREGATE EFFICIENCY BUT DISAGGREGATE INEFFICIENCY

By M.H. ten Raa

November 2003

ISSN 0924-7815

**Bob Russell Volume**

# Don’t Aggregate Efficiency But Disaggregate Inefficiency

**October 2003**

**Thijs ten Raa**

Tilburg University
Box 90153
5000 LE Tilburg
the Netherlands
tenRaa@UvT.nl
Tel +31-20-6252240
Fax +31-20-4206502

## 1. Introduction

Recently Bob Russell published an impossibility result with Chuck Blackorby. The gentlemen argue that it is impossible to aggregate efficiency indices. Since some of us make a descent living decomposing the efficiency of an economy into sectoral contributions the question raises if we are crooks. This paper attempts to give an answer to this question.

Blackorby and Russell (1999) state:

"Perhaps more disturbing is the fact that the principal indexes proposed by Debreu (1951) and Farrell (1957), by Färe and Lovell (1978) and by Zieschang (1983) cannot satisfy these [aggregation] conditions for any technologies, even linear ones."

The subsequent relaxation of these aggregation conditions by Blackorby and Russell (1999) offers little comfort:

"In particular, the [relaxed] aggregation condition provides a rationalization of the Debreu/Farrell efficiency measure, albeit for a very restrictive (linear) class of technologies."

A first step in the process of recovery from bad news is to take stock of the issues. I shall clarify a number of things. First, what are efficiency measures? The Debreu/Farrell name, however much in the air as a reference for a general measure of efficiency, is misleading. The Debreu and Farrell measures better be delineated vis-à-vis each other. Second, can we disaggregate these measures?

The upshot of this paper is a redirection of the measurement of efficiency: top-down instead of bottom-up. I hope the reader will feel better again.

## 2. Efficiency measures: Debreu and Farrell or Diewert?

Consider an economy comprising $l$ commodities, $m$ consumers, with preference relationships $\succsim_i$ and observed consumption vectors $\mathbf{x}^0_i \in \mathbb{R}^l$ ($i$ = 1, ..., $m$), and $n$ production units with sets of possible (net) input vectors $Y_j \subset \mathbb{R}^l$ containing the observed input vector $\mathbf{y}^0_j$ ($j$ = 1, ..., $n$). A combination of consumption vectors and an input vector is *feasible* if the total sum does not exceed the vector of *utilizable physical resources*, $\mathbf{z}^0$, which is the datum of the economy. This constraint is binding for the observed inputs:

$$\sum \mathbf{x}^0_i + \sum \mathbf{y}^0_j = \mathbf{z}^0 \tag{1}$$

The *better set* of net consumption vectors is defined by

$$\mathcal{B} = \{\sum \mathbf{x}_i \mid \mathbf{x}_i \succsim_i \mathbf{x}^0_i, i=1,...,m\} + \sum Y_j \tag{2}$$

Debreu (1951) defines the *coefficient of resource allocation* by

$$\rho = \text{Max } \mathbf{p}(\mathbf{z})\cdot\mathbf{z}/\mathbf{p}(\mathbf{z})\cdot\mathbf{z}^0 \text{ subject to } \mathbf{z} \in \mathcal{B}^{\min} \tag{3}$$

Coefficient $\rho$ measures the distance from the set of minimally required physical resources, $\mathbf{z} \in \mathcal{B}^{\min}$, to the utilizable physical resources, $\mathbf{z}^0$, in the metric of the supporting prices (which indicate welfare indeed). Debreu (1951, p. 284) proves that the distance or the Max in (3) is attained by

$$\mathbf{z} = \rho\mathbf{z}^0 \in \mathcal{B}^{\min}$$

(4)

In modern terminology, this result means that $\rho$ is the *input-distance function*, determined by the program

$$\text{Min } \rho \text{ subject to } \sum \mathbf{x}_i + \sum \mathbf{y}_j \leqq \rho \mathbf{z}^0,\ \mathbf{x}_i \succsim_i \mathbf{x}^0{}_i,\ \mathbf{y}_j \in Y_j \tag{5}$$

Farrell (1957) decomposes efficiency in technical efficiency and allocative efficiency. He notes the similarity between his technical efficiency measure and the Debreu coefficient of resource utilization. Indeed, both concepts are defined through proportionate input contractions. Nonetheless, the analogy is sheer formality and confusing at a conceptual level. It suggests that Farrell takes the Debreu coefficient, augments it, and thus constructs a more encompassing overall measure. It is the other way round; the sway of the Debreu coefficient is far greater than that of Farrell's measure. Particularly Farrell's allocative efficiency measure is a partial (dis) equilibrium concept, conditioned on prices. It takes into account the cost reduction attainable by changing the mix of the inputs, *given* the prices of the latter. The Debreu coefficient, however, is a general (dis)equilibrium concept. It measures the technical and allocative inefficiency in the economy given only its fundamentals: resources, technology, and preferences. Prices are derived and enter the definition of the Debreu coefficient, see (3). Debreu *proves* that the coefficient can be freed from these prices, by formula (4) or nonlinear program (5). Prices remain implicit, however. They support the better set in the point of minimally required physical resources and will be revealed in this paper. The Debreu coefficient measures technical and allocative inefficiency, both in production and consumption, solving the formidable difficulty involved in assessing prices, referred to by Charnes, Cooper, and Rhodes (1978, p. 438). Farrell refrains from this, restricting himself to technical efficiency and price-conditioned allocative efficiency.

The formal analogy between the Debreu coefficient and the Farrell measure of technical efficiency prompted Färe and Lovell (1978) to coin the phrase "Debreu-Farrell measures of efficiency." This is confusing. Debreu's coefficient of resource allocation encompasses both Farrell's technical efficiency and his allocative efficiency measures, plus frees the latter from prices. On top of this, Debreu's coefficient also captures consumers' inefficiencies. The confusion persists. In a very recent review of Farrell's contribution Førsund and Sarafoglou (2002, footnote 4) state

"(Debreu) worked only from the resource cost side, defining his coefficient as the ratio between minimised resource costs of obtaining a given consumption bundle and actual costs, for given prices and a proportional contraction of resources."

Debreu (1951) calculates the resource costs *not* of a given consumption bundle, but of an (intelligently chosen) Pareto equivalent allocation. (And the prices are *not* given, but support the allocation.)

Yet, let me bridge the difference. Following Diewert (1983), I limit inefficiency to production by assuming Leontief preferences. Under this assumption Debreu's program (5) can be shown to reduce to

$$\text{Min } \rho \text{ subject to } \sum \mathbf{x}^0_i + \sum \mathbf{y}_j \leqq \rho \mathbf{z}^0,\ \mathbf{y}_j \in Y_j \tag{6}$$

The detail is in ten Raa (2003) who calls the consequent $\rho$ the *Debreu-Diewert efficiency measure.*

## 3. An example

Let us consider an economy producing a single consumption good. Denote the inputs by vector $\mathbf{l}$. The available stock of inputs is $\mathbf{l}^0$. The production possibilities are given by two production functions, one for each unit: $F_1$ and $F_2$. The observed inputs are $\mathbf{l}^0_1$ and $\mathbf{l}^0_2$. Efficiency program (6) reads

$$\text{Min } \rho \text{ subject to } \sum \mathbf{x}^0_i \leqq F_1(\mathbf{l}_1) + F_2(\mathbf{l}_2) \text{ and } \mathbf{l}_1 + \mathbf{l}_2 \leqq \rho \mathbf{l}^0 \tag{7}$$

The solution denotes the Debreu-Diewert efficiency of the economy. Denote the efficient inputs by $\mathbf{l}_1$ and $\mathbf{l}_2$ and contrast them with the observed inputs. The question is: How efficient are the units? It may very well be that both units produce the maximum output given their inputs, but that the distribution of inputs is inefficient. For example, if there is only one input and the units have the same, strictly concave production function, say $F$, then the efficient distribution of inputs is fifty/fifty. This example, however simple, conveys the message of Blackorby and Russell (1999). The efficiency of the units in the sense of maximizing output given the inputs does *not* imply that the constellation of the two units is efficient.

Is there no way to cope with this example? My idea is to look at profits, not at market prices, but at shadow prices. Choosing the consumption good as numerair, the shadow prices of the inputs are their marginal products or the vector of partial derivatives $F'(\rho\mathbf{l}^0/2)$, evaluated at the optimum. These input prices will be intermediate, higher than the marginal product of a big unit, smaller than the marginal product of a small unit (assuming concavity). I shall consider the small unit relatively efficient. The big unit will pick up more inefficiency.

## 4. Back to the model

Let the production possibility set be given by $Y_j = \{\mathbf{y}_j | F_j(\mathbf{y}_j) \geq 0\}$ where the differentiable functions $F_j$ are concave. (In the previous example these functions map $(\mathbf{l}, -x)$ into $F_j(\mathbf{l}) - x$: the value of the production function of the previous section at input vector $\mathbf{l}$ minus output. Since the functions are concave the differentiability assumption can be dropped and the subsequent analysis would be in terms of subgradients.) Efficiency program (6) reads

$$\text{Min } \rho \text{ subject to } \sum \mathbf{x}^0_i + \sum \mathbf{y}_j \leqq \rho \mathbf{z}^0,\ F_j(\mathbf{y}_j) \geq 0 \tag{8}$$

Unlike the Blackorby and Russell (1999) condition, $F_j$ need not be linear. Consequently, (8) is a nonlinear program. According to Wolfe (1961) the dual program is

$$\text{Max}_{\rho,\,\mathbf{y},\,\mathbf{p},\,\boldsymbol{\tau}}\ \rho - \mathbf{p}(\rho\mathbf{z}^0 - \textstyle\sum\mathbf{y}_j - \sum\mathbf{x}^0{}_i) - \boldsymbol{\tau}\mathbf{F}(\mathbf{y})$$
$$\text{subject to } \mathbf{p}\mathbf{z}^0 = 1,\ \mathbf{p} = \boldsymbol{\tau}\mathbf{F}'(\mathbf{y}),\ \text{and } \mathbf{p},\boldsymbol{\tau} \geq \mathbf{0} \tag{9}$$

and by his Theorem 2 (9) has the same solution value as (8). Here **F** is the vector with components $F_j$ and **F**' is the matrix with the *j*-th row displaying the partial derivatives of $F_j$. Notice that the first two terms in (9) cancel by the first dual constraint.

The analysis becomes highly transparent if we assume constant returns to scale, or, in Blackorby and Russell (1999) jargon, linear homogeneity. In this case Blackorby and Russell (1999) argue that the inputs must be perfect substitutes and the outputs must be perfect substitutes. This prohibitive restriction rules out CES functions and even the case of fixed input coefficients, and reduces (9) to a linear program. Under general linear homogeneity, however, (9) remains a nonlinear program, but the third and fifth terms in (9) cancel by the second dual constraint and Euler's theorem. Hence only the fourth term remains and we obtain

$$\text{Max}_{\mathbf{y},\,\mathbf{p},\,\boldsymbol{\tau}}\ \mathbf{p}\textstyle\sum\mathbf{x}^0{}_i \text{ subject to } \mathbf{p}\mathbf{z}^0 = 1,\ \mathbf{p} = \boldsymbol{\tau}\mathbf{F}'(\mathbf{y}),\ \text{and } \mathbf{p},\boldsymbol{\tau} \geq \mathbf{0} \tag{10}$$

And, by Wolfe's Theorem 2,

$$\rho = \mathbf{p}\textstyle\sum\mathbf{x}^0{}_i \tag{11}$$

Now value the *observed* inputs and resources, see equation (1), by the *shadow* prices. Substituting (10) and (11), and rearranging terms, I obtain

$$1 - \rho = \sum \mathbf{p}\, \mathbf{y}^0_j \tag{12}$$

On the left hand side is inefficiency and on the right hand side are losses at shadow prices. (Remember $\mathbf{y}^0_j$ are net input vectors.) The shadow prices are given by the second dual constraint of (9) or (10), namely $\mathbf{p} = \boldsymbol{\tau}\mathbf{F}'(\mathbf{y})$. These are the marginal products of the efficient units. If a unit is inefficient, that is within its own frontier—$F_j(\mathbf{y}_j) > 0$—then $\tau_j = 0$ by the phenomenon of complementary slackness (which is equivalent to Wolfe's Theorem 2) and it plays no role in price formation.

## 5. Another example

Consider an economy with one input ($L$) and one output ($Y$). The production possibilities for two units (1 and 2) are $Y \leq L$ and $Y \leq \beta L$, respectively. The observed allocation is $(Y^0{}_1, L^0{}_1) = (½, ½)$, $(Y^0{}_2, L^0{}_2) = (½\beta, ½)$. Notice that both units are efficient in the sense of being on their frontiers. Blackorby and Russell (1999) argue that output (input) aggregation of efficiency indices is possible only if the efficiency indices are ratios of linear functions of input and output quantities and the aggregate index is a weighted average. Moreover, these functions must be *common* to all units. This implies the first of the following statements.

I. *The economy is efficient if and only if* $\beta = 1$.

II. If $\beta < 1$, the efficient allocation is
$(Y_1, L_1) = (½ + ½\beta, ½ + ½\beta)$, $(Y_2, L_2) = (0, 0)$.
Hence efficiency $\rho = ½ + ½\beta$ and inefficiency $1 - \rho = ½ - ½\beta$.

III. If $\beta > 1$, the efficient allocation is
$(Y_1, L_1) = (0, 0)$, $(Y_2, L_2) = (½ + ½\beta, ½\beta^{-1} + ½)$.
Hence efficiency $\rho = ½\beta^{-1} + ½$ and inefficiency $1 - \rho = ½ - ½\beta^{-1}$.

To decompose inefficiency, let us first determine the shadow prices. The first dual constraint, see (9) or (10) normalizes the shadow price of the input to 1. The shadow price of the output is 1 in cases I and II and $\beta^{-1}$ in case III.

I. Both units make zero losses at shadow prices, hence pick up zero inefficiency. This is in perfect agreement with Blackorby and Russell (1999).

II. At shadow prices the losses of the units are ½ – ½ and ½ – ½$\beta$, respectively. The inefficiency is imputed entirely to the second unit. Indeed, it should be out of business.

III. At shadow prices the losses of the units are ½ – ½$\beta^{-1}$ and ½ – ½$\beta\beta^{-1}$, respectively. The inefficiency is imputed entirely to the first unit. Indeed, it should be out of business.

The point of the example is that inefficiency has been decomposed in cases where efficiency cannot be aggregated according to Blackorby and Russell (1999). The result holds for *nonlinear* technologies, like CES functions, including the limiting case of a Leontief function.

## 6. Conclusion

To determine the efficiency of a constellation of production units we need the following data.

a. The inputs and outputs of each unit
b. The production possibilities of each unit

Notice that this is no more than what is required by Blackorby and Russell (1999).

I suggest we proceed as follows. The first step is to compute the efficiency of the system of the units. This is done by contracting the total input of the system subject to the condition that total output is preserved, allowing for reallocations of the inputs and outputs between the units. The percentage by which contraction is feasible is the inefficiency in the economy. The second step is to compute the shadow prices of the contraction program. They are the marginal products of the efficient units. The output shadow prices will be low, as they reflect best practice costs. The third step is to value the units (in terms of profits) at shadow prices. Under constant returns to scale the best practice units break even; their values are zero. The other units incur losses though. This paper has shown that *the losses sum to the aggregate inefficiency*. This completes the decomposition of inefficiency.

The inefficiency of a unit can have *two* sources. First, the unit may operate within its possibility frontier. Second, the unit may produce the wrong output vector, not the one implied by the optimal allocation of inputs between units. *This allocative source of firm inefficiency should not be neglected.* If one neglects it, one obtains impossibility results on the aggregation of efficiency indices. If one takes it into account, inefficiency *can* be disaggregated. With Richard Nixon, let me conclude "I am not a crook."

## 7. References


Charles Blackorby and R. Robert Russell (1999) "Aggregation of Efficiency Indexes," *Journal of Productivity Analysis* 12, 5-20

Charnes, A., W.W. Cooper, and E. Rhodes (1978) "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research* 2, 419-444.

Debreu, Gerard (1951) "The Coefficient of Resource Utilization," *Econometrica* 19, 3, 273-292.

Diewert, W. Erwin (1983) "The Measurement of Waste Within the Production Sector of an Open Economy," *Scandinavian Journal of Economics* 85, 2, 159-179.

Färe, Rolf and C.A. Knox Lovell (1978) "Measuring the Technical Efficiency," *Journal of Economic Theory* 19, 1, 150-162.

Farrell, M.J. (1957) "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society* Series A, 120, III, 253-281.

Førsund, Finn R. and Nikias Sarafoglou (2002) "On the Origins of Data Envelopment Analysis," *Journal of Productivity Analysis* 17, 23-40.

Hulten, Charles R. (1978) "Growth Accounting with Intermediate Inputs," *Review of Economic Studies* 45, 3, 511-518.

Jorgenson, Dale W., and Zvi Griliches (1967) "The Explanation of Productivity Change," *Review of Economic Studies* 34, 3, 308-350.

Nishimizu, Mieko, and John M. Page, Jr. (1982) "Total Factor Productivity Growth, Technological Progress and Technical Efficiency Change: Dimensions of Productivity Change in Yugoslavia," *Economic Journal* 92, 920-936.

ten Raa, Thijs (1995) *Linear Analysis of Competitive Economies* LSE Handbooks in Economics. Hemel Hempstead: Harvester Wheatsheaf/Prentice Hall.

Wolfe, Philip (1961) "A Duality Theorem for Nonlinear Programming," *Quarterly of Applied Mathematics* 19, 3, 239-244.

Zieschang, K. (1983) "An Extended Farrell Technical Efficiency Measure," *Journal of Economic Theory* 33, 387-396.